DATA GOVERNANCE FRAMEWORK

The Data Governance Framework (DGF) has been designed with the vision of creating an Integrated Data System in Pakistan

1. Vision and objectives of the Data Governance Framework (DGF)

The Data Governance Framework (DGF) has been designed with the vision of creating an Integrated Data System in Pakistan. An effective Data Governance Framework is crucial not only for regulatory compliance but also for enhancing operational efficiency and cultivating public trust in official statistics. In this context, the DGF aims to:

1.1. Serve as a strategic framework to align and harmonise the operations of the National Statistical System (NSS), particularly the Pakistan Bureau of Statistics (PBS) and the provincial bureaus of statistics (PBoS).

1.2. Establish overarching guidelines for standardised procedures and protocols for data collection, protection, coordination, and dissemination to ensure consistency and comparability across all provincial bureaus of statistics.

1.3. Strengthen the capacity and efficiency of the 'bureaus of statistics' to establish an effective data management system.

2. Scope of the Data Governance Framework (DGF)

2.1. *The Data Governance Framework* (hereafter, the DGF) acts as an all-encompassing guide aimed at enhancing the day-to-day functional and operational capacities of the bureaus of statistics.

2.2. The DGF, primarily designed for PBS and the PBoS, but applicable to any statistical organisation, outlines an organised framework of policies, procedures, standards, and tools to ensure high quality, consistency, security, and ethical data usage across all operations.

2.3. The DGF seeks to establish a coordination mechanism among departments while promoting the standardisation and harmony of data processes across all bureaus of statistics.

2.4. By outlining the mechanism and standard operating procedures (SOPs) for essential day-to-day operations—such as data gathering, dissemination, coordination, and protection—the DGF offers an integrated framework that promotes and standardises efficient data management and dissemination.

2.5. The DGF operates within the established framework of the *General Statistics (Reorganization) Act, 2011*. It functions as a practical tool, delivering a clear layout of SOPs that can be implemented to facilitate operational efficiency.

2.6. Since the DGF does not propose any changes to existing laws or policies governing National Data Systems, its implementation does not necessitate legal approvals. Approvals from the heading authority of the organisation, primarily bureaus of statistics in the current context, shall be sufficient.

3. Landscape of data governance in Pakistan

3.1. The National Statistical System (NSS) of the country, consisting of the federal and provincial bureaus of statistics (PBS and PBoS respectively hereafter), government departments/ministries, and the State Bank of Pakistan, is primarily accountable for data collection, protection, and dissemination.

3.2. The General Statistics (Reorganization) Act, 2011, provides legal cover and mandates for the NSS but does not constitute any guidelines for the day-to-day operations of the entities that fall under the NSS.

3.3. There are currently no established guidelines on data collection, privacy and security, management, dissemination, stakeholder collaboration or the standardisation of data collection and reporting procedures across different departments, especially within provincial bureaus.

3.4. In the absence of unified operational guidelines, each provincial bureau independently follows its own practices for data collection, processing, management, and dissemination.

3.5. Functions like acquiring data across various departments are conducted based on traditional and personal relations, instead of a formalised system.

3.6. Provinces lack policies or frameworks for data dissemination, protection and coordination. Therefore, the practices to collect, disseminate, protect and coordinate data vary by data producers, particularly the bureaus of statistics.

4. The need for the DGF

The critical role of data in modern development planning, Pakistan's renewed focus on evidence-based policymaking, the essential contribution of the PBoS to provincial development planning, and the pressing demand for high-quality and timely data all underscore the need for a DGF to manage data across the country. Consequently, the following factors contribute to the dire need for a DGF:

4.1. Absence of any operational document governing data in the country, especially at the provincial level.

4.2. An ambiguous mechanism and lack of SOPs guiding the day-to-day functions of i) data collection; ii) data protection; iii) data dissemination; and iv) data coordination.

4.3. The need for harmonisation of data standards, data quality and minimising discrepancies across all the bureaus of statistics and the NSS in general.

4.4. Streamlining data coordination within and across various data producers, the PBS and PBoS, and government departments and data users.

4.5. Establishing institutional data coordination between the PBS and PBoS and distinct departments. For instance, when a bureau requires data from a specific department, that department is not obligated to fulfil the request.

4.6. Reducing the bottlenecks in departmental/secondary data collection—a major job of provincial bureaus—by introducing standard mechanisms and SOPs. At present, there is no binding principle that provides a specific timeline for data sharing, severely limiting the data coordination between the bureaus and departments and between the PBS and the bureaus.

4.7. Enhancing usage of data through a standardised data dissemination strategy, which is common across the NSS, especially federal and provincial bureaus of statistics. Currently, there is no uniform provincial data dissemination policy in place[1].

4.8. Providing guidelines for data protection, which are robust, updated and standardised across different data producers, specifically the PBS and PBoS.

[1] The first national data dissemination policy made by the Pakistan Bureau of Statistics is currently under process.

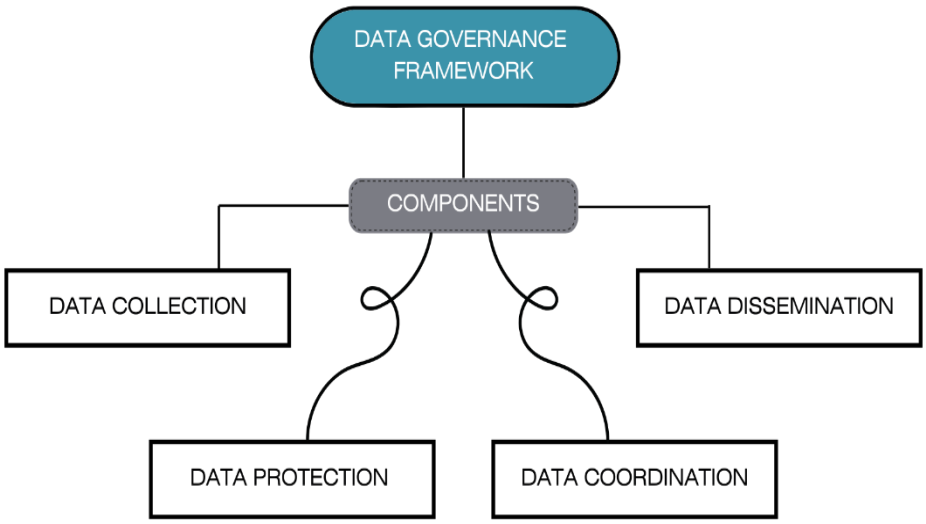
Currently, there are no guidelines present on how to collect, store, and protect data or even retrieve lost data. These factors culminate as barriers to the quality and availability of data whilst presenting further challenges within the bureaus.

The DGF addresses these gaps by providing a comprehensive framework that outlines uniform mechanisms and SOPs to ensure the quality of data—including reliability, coherence, accessibility, clarity, and comparability[2]. This framework enables improved collaboration, harmony and standardisation of functions and data across national and provincial bureaus and departments.

5. Fundamental components of the DGF

The DGF adopts an all-encompassing approach that covers the entire data lifecycle—ranging from data collection, processing and protection to dissemination. The DGF serves as an overarching strategic guide, providing a comprehensive and adaptable structure for data governance practices. Its primary function is to establish a unified direction and set of principles for managing data across various provincial departments and agencies. The DGF encompasses four key components as shown in Figure 2.

Figure 2: Components of a Data Governance Framework



5.1.1. Data collection: This forms the foundation of any data governance strategy and the DGF. It involves the systematic gathering of data from various sources, ensuring that the data is accurate, reliable, and relevant. The DGF provides guidelines which ensure standardisation, consistency and quality of data collection processes and methods across multiple sectors.

5.1.2. Data protection: Protecting data from breaches, misuse and leakages is highly important. The DGF outlines SOPs for this very purpose, including the

[2]https://unstats.un.org/UNSDWebsite/nqaf/Attachment%20B-UNNAQF%20Manual%20in%20PDF.pdf

implementation of robust security measures, encryption, access controls, and regular audits.

5.1.3. Data coordination: Effective data governance requires seamless coordination among all relevant stakeholders, including government departments, private organisations, civil society organisations, and academia. Effective coordination SOPs stated in the DGF ensure that data is shared, integrated and utilised effectively and ethically, avoiding duplication and fostering collaboration.

5.1.4. Data dissemination: Disseminating data to relevant users in a timely and accessible manner is crucial for informed decision-making. This involves developing policies and platforms for data sharing, ensuring data is available to all authorised users.

6. Founding principles of the DGF

Achieving robust and effective data governance requires a well-defined set of guiding principles to inform strategies, policies and implementation. Adherence to these principles enhances the quality and credibility of official statistics, enabling effective decision-making, fostering operational efficiency, and facilitating the development of high-standard service delivery policies.[3] The following principles form the fundamentals of the DGF:

6.1. Data accuracy is fundamental for effective governance. Ensuring accurate, valid and reliable data is crucial for informed decision-making, operational efficiency, and service delivery. This principle involves strategies to maintain data accuracy throughout its lifecycle, including data validation, cleaning, and regular audits.

6.2. Data accessibility ensures timely, reliable, and secure data access across all authorised users. A robust data governance framework guarantees that the right individuals have access to the necessary data, balancing accessibility and data privacy and security. This involves managing roles and access controls alongside the assurance of secure data storage and transmission.

6.3. Data consistency safeguards uniform data definitions, formats, and values across systems and processes. Using uniform data models and employing data integration tools are essential to establishing and enforcing these standards. Procedures should be established to effectively address and resolve discrepancies.

6.4. Data handling and management shall be compliant with the legislation and policies governing data systems—Act of 2011, in Pakistan's context. Adherence to relevant data laws, regulations, and standards is essential for effective data governance. A comprehensive governance framework includes policies addressing data collection, processing, storage, and dissemination. Regular audits and compliance checks reinforce adherence, while employee training embeds a culture of compliance within the organisation.

6.5. Data integrity guarantees that information remains accurate and reliable throughout its lifecycle. Measures such as data validation, backups, security controls, and version control mechanisms safeguard against data loss and unauthorised modifications, supporting apt decision-making and thorough compliance.

---

[3] https://datagalaxy.com/en/blog/7-core-principles-of-data-governance/#:~:text=Data%20stewardship%20is%20one%20of,accessible%2C%20consistent%2C%20and%20protected

6.6. Data stewardship encompasses the administration and oversight of data assets, ensuring they are high-quality, consistent, protected, and accessible. Effective stewardship requires a clear assignment of roles and responsibilities for data management tasks, including maintaining data quality and ensuring compliance with established standards and regulations.

6.7. Data transparency encompasses a commitment to openness regarding the processes of data collection, storage, management, and usage processes. It fosters trust among stakeholders and compliance per data protection laws. Clear policies, regular communication about such practices, and access to vital information on data processes are essential for demonstrating accountability and ethical practice.

7. Mechanism and SOPs for data collection

7.1. Objectives

7.1.1. Accuracy and reliability:

Ensure the accuracy and reliability of data through processes such as implementing standardised procedures, rigorous quality controls, and validation of data collection.

7.1.2. Comprehensive coverage:

Achieve comprehensive coverage of various sectors such as health, education, agriculture, economy, and demographics to provide a holistic view of the province's standing.

7.1.3. Timeliness:

Guarantee the prompt collection and processing of data to support real-time decision-making and policy formulation.

7.1.4. Consistency and comparability:

Maintain consistency in data collection definitions and methods over time to ensure comparability across different periods and regions.

7.1.5. Transparency and accountability:

Promote transparency and accountability in the data collection process by clearly documenting methodologies, sources, and any adjustments made to the data, with information being publicly available.

7.1.6. Data security and confidentiality:

Implement robust data security measures to preserve the confidentiality of individuals and sensitive information.

7.1.7. Capacity building:

Enhance the capacity of data collection and analysis personnel and departments through regular training and development programs, advancing the quality and efficiency of data collection. A standard annual training calendar, consistent across all provinces, shall be set in place.

7.1.8. Stakeholder engagement:

Engage with stakeholders—including government agencies, academia and the private sector—to ensure the data collected meets their requirements and supports their broader development goals.

7.2. Guiding principles for data collection[4]

Data collection shall:

i) Have standardised mechanisms and SOPs across the PBS and PBoS
ii) Facilitate the generation of relevant, high-quality and timely statistics, guaranteeing accuracy and reliability
iii) Respect respondents' time and resources, applying uniform procedures across all activities
iv) Strive to earn the trust of information providers by respecting respondents' privacy and maintaining the confidentiality of their data, while promoting transparency and upholding ethical standards.
v) Be cost-effective in resource utilisation, collecting data promptly to maintain its relevance.

7.3. Mechanisms for efficient data collection

7.3.1. Identify data needs:

Conducting a data need assessment is the first step in the data collection process. This involves collaborative efforts with government institutions, NGOs, and other relevant stakeholders to acquire the specific data needed for policymaking, programme implementation, and research. Data needs assessments shall be conducted once every three years, with a maximum interval of five years.

7.3.2. Update data collection methods:

Survey instruments shall be updated based on findings from the data needs assessment. This process may include adding new sections or new questions, rephrasing existing questions, or revising response options to accommodate new respondent groups or other emerging requirements. Such updates would involve (re)designing questionnaires, updating data collection protocols and improving data storage facilities. The primary objective would be to ensure that the data obtained is accurate, valid, and credible across different regions and time periods. All PBoS should establish dedicated survey instrument updating bodies, which should meet at least once a year.

7.3.3. Implement data collection:

Implementing the data collection involves deploying trained personnel to collect primary data from respondents and secondary data from relevant sources. It is important to verify that the data collection process is conducted in a timely and efficient manner, with strict adherence to designed methods to maintain data quality and integrity. Digital dashboards, based on survey CTOs, shall be established in all bureaus to monitor and analyse all the data collection—including hardcopy data.

7.3.4. Ensure quality control:

Mechanisms for quality control for reviewing and confirming data involve conducting data audits, cross-checking data from different sources, and validating data consistency and accuracy. It is crucial to maintain high standards of data quality to ensure that the data collected is reliable and promptly applicable for decision-making and policy formulation in Pakistan.

---

[4] https://nsa.nsa.org.na/wp-content/uploads/2021/04/Data-Collection-Processing-and-Disemination-Policy.pdf

7.4. Standard Operating Procedures **(**SOPs)

7.4.1. Planning and preparation

- Clearly define the objectives of the data collection process in alignment with provincial development goals.
- Adopt standardised sampling techniques (random, stratified, or cluster sampling) to ensure representative and unbiased data.
- Use pre-tested, structured questionnaires or digital tools to ensure consistency in data collection across various sectors.

7.4.2. Sustainable training system for enumerators and bureau staff

- Ensure comprehensive training for supervisors and field enumerators on the data collection process, emphasising technical skills and ethical guidelines.
- An annual training calendar, consistent across all provinces and clearly marking the periods allocated for training, shall be set at the beginning of each year. The calendar shall be publicly available so that the government, development partners and other stakeholders can align their training schedule/programme accordingly.
- Undertake periodic (maximum every third year), capacity gaps/training needs assessments of the departments and staff. These assessments shall be shared with Planning and Development Departments (P&DDs) and relevant partner organisations.
- A comprehensive training database shall be developed to record and track details of all participants, including the form of training attended, duration of each session, and relevant participant information. This database would serve as a central repository to monitor training progress, identify capacity-building needs, and ensure efficient management for future training programmes.

7.4.3. Data collection procedures

- As per feasibility, integrate electronic data collection tools (such as tablets or mobile applications) for real-time data entry, minimising errors and delays.
- Conduct pilot testing of data collection tools and processes in small areas to identify issues and improve data accuracy before full-scale implementation.
- Ensure data confidentiality, in accordance with national and international data privacy regulations.

7.4.4. Integration of new technologies and data collection methods

- Expansion of roles from data collectors to data custodians by collecting existing data from third-party resources, such as the internet and private-sector databases.
- Utilising advanced technologies and sources for data collection to minimise existing limitations and costs. For instance, verified third-party providers like Google and Microsoft can supply geospatial data, offering insight into sectors such as agriculture and transport. This can be combined with other data sources such as the census[5].
- When obtaining data from third-party sources, including the internet, strict attention shall be paid to the quality and credibility of data, including methodology, sources and purpose of collection.

[5] https://www.oecd-ilibrary.org/sites/276aaca8-en/1/2/5/index.html?itemId=/content/publication/276aaca8-en&_csp_

- Provide IT training and knowledge where requisite.

7.4.5. Field operations (for primary data collection)

- Assign field supervisors to monitor the progress of enumerators, ensuring adherence to procedures and addressing on-ground challenges immediately.
- Field teams shall submit daily progress reports to supervisors, stating any irregularities or difficulties encountered during data collection.
- Implement random spot-checks by supervisors and cross-verification of collected data for accuracy and consistency.
- Establish live data dashboards, with survey CTOs, to monitor real-time data collection from the field. All the interviews shall be recorded on a survey CTO. A daily report analysing the quality of data collected shall be published and examined each day by an expert.

7.4.6. Data management

- After collection, data shall be promptly entered into the digital system and validated for errors. Use software tools (e.g., Stata, SPSS) to clean the data and remove inconsistencies.
- Ensure that all collected data is backed up in secure, encrypted databases, with regular backups maintained to avoid loss.

7.4.7. Data quality and transparency checks

- PBoS and data collection authorities shall conduct initial checks for completeness and accuracy upon receiving data submissions. Any discrepancies or incompletions shall be flagged within three working days, with departments given one week to resolve the issues.
- Ensure a transparent data collection process (e.g., in the case of surveys, a detailed methodology should be compiled and made easily available).
- In the case of collecting individuals' data, the participants shall be informed of what data is being collected, how it will be stored, who will have access to it, and how it may be used.

8. Mechanisms and SOPs for data protection

Data protection involves measures to safeguard data from unauthorised access, storage, use, disclosure, modification or destruction, ensuring that the data collected is securely protected and maintained. This essentially guarantees that data storage adheres to the set regulations and laws.

8.1. Objectives

i) Protect data from unauthorised access, breaches, and storage.
ii) Ensure data remains accurate and unaltered.
iii) Adhere to all relevant data protection laws and regulations.
iv) Implement robust security protocols to protect data.
v) Build confidence among stakeholders through effective data protection measures.

8.2. Mechanisms

8.2.1. Develop data protection policies: The PBS and PBoS shall establish policies that define the standards for protecting data, including specifying access permissions, storage locations, and management protocols, to ensure its security.

8.2.2. Ensure privacy compliance: Implement practices to ensure that data collection, storage, and usage comply with national and provincial privacy laws and regulations. These measures aim to protect informants' privacy rights.

8.2.3. Backup and recovery: Develop and implement robust data backup and disaster recovery plans to protect data against loss or damage. This would ensure data availability, continuity, and quick retrieval in case of an emergency.

8.3. Data privacy and security

Unrestricted and unadjusted information about individuals raises privacy and data security concerns throughout the data lifecycle. To address this, two main policy levers have been proposed. First, the PBoS will classify data aggregation, availability, and access based on its nature. Second, individuals will have rights over their data.

8.3.1. Aggregation, anonymity and access:

Unless a legitimate purpose is demonstrated, and the data user assumes heightened responsibility and legal culpability for any misuse, identifying information and sensitive characteristics—like religion, race/caste, location, or address—shall be removed or anonymised in publicly shared data. However, advances in computing can still pose a risk of re-identification. To further protect anonymity while preserving usability, data should be aggregated as much as possible without compromising the user's ability to derive meaningful insights. Decisions to withhold or release data, balancing potential risks and benefits, fall to the PBoS.

The PBoS will categorise data into two types: public-facing (Public Use Files – PUF) and licence-restricted (Micro Data Under Contract – MUC). Public-facing data is highly anonymised, pooling data and removing individual details before public dissemination. Licence-restricted data, with more geographic and variable details, requires additional safeguards, such as user restrictions and contracts assigning responsibility for data misuse to the user.

8.3.2. Data subject rights:

Individuals questioned during data collection shall have the right to withdraw consent and request data deletion or correction to maintain the integrity of state records.

8.4. Standard Operating Procedures

8.4.1. Data access control and authorization

- Assign data access rights according to predefined roles to ensure that only relevant personnel have access to certain datasets or statistical systems.
- Conduct regular audits (at least quarterly) of user access permissions to confirm compliance with internal policies.
- Revoke access for employees who no longer need it due to role changes or termination.

8.4.2. Data encryption and secure storage

- Use robust encryption protocols for data in transit across networks.
- All backups will also be encrypted.
- Sensitive data, including databases and file servers, shall be stored in encrypted environments to prevent unauthorised access.
- Encrypt backup files and protect them with strong access controls to ensure that backups are regular and stored in secure, off-site locations.
- Each province will have its own data resource centre for secure storage of data.

8.4.3. Data retention and disposal

- Establish a data retention policy that specifies the duration of different types of data (e.g., survey data, administrative data) to be stored before disposal. Retention periods should comply with national laws and organisational needs.
- Implement automated data retention schedules that ensure data is securely archived or deleted after the retention period has expired.
- Use secure methods for data disposal, such as permanent deletion (data wiping) of electronic records and shredding of paper records.
- Ensure all electronic devices (e.g., hard drives) are securely cleaned or physically destroyed before disposal.

8.4.4. Data sharing and secure transfer

- Before sharing data, verify that formal *Data Sharing Agreements* are signed/filled, outlining the purpose, scope, and security obligations of both parties.
- Use secure transfer methods when transmitting sensitive data. Avoid using unsecured channels like email for sensitive data transfers.
- Ensure that only authorised recipients have access to the shared data by using encryption keys or temporary access credentials.
- Provide employees with clear guidelines and best practices on securing workstations, managing passwords, and safeguarding sensitive data, both in-office and remotely.

8.4.5. Monitoring and evaluation

- Conduct regular security audits (at least annually) of all IT systems, data handling practices, and employee access controls to identify potential vulnerabilities.

9. Mechanisms and SOPs for data coordination

The Bureau of Statistics facilitates data coordination for collection and dissemination through two primary channels: coordination between the PBS and PBoS, and coordination between the PBoS and other departments. Over two-thirds of provincial-level data collection involves gathering information from different public-sector departments,[6] making data coordination the most critical function of provincial bureaus.

The mechanisms and SOPs for data coordination should enhance the collection, collaboration, and integration of data among various stakeholders and government departments. Effective coordination among data producers aims to enhance the quality of statistics, support development planning, and ensure the implementation of statistical activities through close

[6] Based on findings from in-depth interviews and consultation with heads PBoS.

collaboration. This, in turn, ensures better resource quality, avoids duplication of efforts, and boosts the public trust in the NSS, thereby increasing user trust in official statistics.[7]

9.1. Objectives

9.1.1. Foster collaboration: The framework provides guidelines to promote partnerships among data producers and stakeholders, improving data sharing, quality, and efficiency to better cater for policymakers and researchers.

9.1.2. Standardised practices: It establishes uniform data collection and sharing protocols to enforce consistency and improve data quality and reliability for decision-making.

9.1.3. Enhance interoperability: The framework seeks to improve the integration of data systems across organisations, facilitating seamless data exchange and efficient management.

9.1.4. Engage stakeholders: Involving all relevant stakeholders ensures inclusivity and enhances the relevance and accuracy of data for policy development.

9.2. Guiding principles

9.2.1. Foster a cooperative environment among stakeholders to enhance data sharing and integration.

9.2.2. Apply uniform standards and protocols for data activities to ensure consistency and compatibility.

9.2.3. Maintain openness in data coordination processes to build trust among stakeholders.

9.2.4. Ensure all relevant stakeholders are involved in data coordination efforts.

9.2.5. Regularly review and improve coordination practices to enhance effectiveness.

9.3. Mechanisms

9.3.1. Establish governance structures: To effectively coordinate data, it is crucial to establish governance structures. A high-powered committee on data coordination shall be established within the National Statistical System (NSS). This committee will oversee data coordination activities, promoting coherence in the development and implementation of statistical programs. According to the United Nations Economic Commission for Europe (UNECE), national coordination bodies play a vital role in ensuring the consistent development and implementation of statistical programmes (UNECE, 2018).

9.3.2. Develop data standards: Standardising data formats, definitions, and protocols is essential for maintaining consistency and comparability in data management. The International Organization for Standardization (ISO) emphasises the significance of data standards in achieving interoperability and consistency in data management (ISO, 2017).

9.3.3. Facilitate collaboration and engage stakeholders: Organising regular meetings, workshops, and forums for stakeholders to collaborate and share insights is key to effective data coordination. This will encourage the crucial aspect of involving relevant stakeholders in data coordination activities, from data collection to dissemination.

[7] https://www.sesric.org/imgs/news/1850-PRESENTATION_NSO_ALBANIA_Maldi%20DEMA_EN.pdf

9.3.4. Establish a forum, National Data Coordination Council/Committee (NDCC), for annual meetings of heads of the PBS and PBoS: A forum shall be established to improve coordination between the heads and management of the PBS and PBoS, complete with the relevant ToRs.

9.3.5. The forum shall comprise the chief statistician of the PBS and the directors-general and directors of the bureaus of statistics from each province. To promote equal participation in leadership positions, the chair of the forum shall rotate after three years, allowing each entity the opportunity to lead the forum. This structure aims to foster balanced representation and collaborative governance with the NDCC.

- The NDCC shall convene annually to discuss and strategise operational procedures aimed at enhancing the data system. Participation from the PBoS would allow greater coordination amongst provincial statistical offices.

9.3.6. Primary functions of the NDCC include:

- Approve annual or quarterly activity plans, outlining strategic initiatives, priorities, and resource allocations.
- Lead efforts to standardise concepts, definitions, methodologies, coverage, classifications, and dissemination practices across the NSS to ensure data consistency and reliability.
- Coordinate main field activities to enhance the standard and productivity of data collection processes.
- Implement best practices and standards to increase the quality of statistics across the NSS.
- Develop capacities within the NSS, ensuring that all data-producing entities have the necessary skills, tools, and knowledge to perform effectively.

9.4. Standard Operating Procedures

9.4.1. Designated Point of Contact (POC)

- There shall be a designated Point of Contact (POC) responsible for handling all data-sharing requests and coordination activities between departments both at the PBoS and line departments.
- The POC will serve as the primary liaison for both internal and external communication related to data sharing.

9.4.2. Stipulated timeframe for initial response to data/other requests

- All departments shall respond to data-sharing requests and coordination queries within a maximum of two weeks.
- Data request system to be readily available online.

9.4.3. Automated email reminders

- An automated system will be implemented, sending reminder emails to relevant departments five days before the deadline for data submission.
- One reminder at five days and a final reminder on the day of the deadline.
- Any delays in response should trigger automatic escalation to senior management for follow-up.

9.4.4. Central data repository

- Have a central repository where data from all government departments, bureaus, and agencies is collected and stored in a standardised format.

- Facilitate the integration of data from various sources (such as health, education, and demographics) into a single repository to avoid duplication.
- Enable remote access to the repository for all bureaus and government departments.
- Develop a metadata catalogue that thoroughly describes the data, enabling stakeholders to easily locate relevant datasets while being prefaced on the context, source, and quality of the data.
- Implement role-based access control to ensure that different degrees of access are granted based on the need and authority of each department.
- Develop an open-access policy, ensuring that departments can securely share their data with other bureaus without compromising sensitive information.
- Data repositories at the individual department/bureau level and granting access to all other stakeholders.
- Access to the repository can be expanded to incorporate the private sector to promote public-private partnerships in data. The private sector can request access to data at the discretion of the relevant bureau or authority.

9.4.5 Annual coordination meetings

- Bi-annual coordination meetings will be scheduled between the PBoS and relevant departments to discuss ongoing data-sharing activities, resolve any coordination gaps, and review compliance with the established SOPs.
- These meetings will serve as a platform for continuous improvement in data management and sharing.

9.4.6. Data validation checks

- The PBoS shall implement data validation checks both before and after sharing data to ensure accuracy, completeness, and consistency.
- These checks should include automated and manual reviews to identify any discrepancies or errors before the data is used or shared with other departments and the public.

9.4.7. Formal data request submission

- All departments shall be required to submit a 'formal data request form' when seeking data. This form should specify the type of data requested, the purpose of the request, the intended use, and the timeline for receiving the data.

10. Mechanisms and protocols for data dissemination

Data Dissemination is a critical component of efficient data governance. This framework aims to ensure that the data collected, processed and analysed by statistical bureaus is effectively forwarded and accessed by relevant stakeholders. By establishing clear guidelines and procedures for data dissemination, the framework seeks to foster transparency, encourage data-driven decision-making, and cultivate public trust in statistical processes.

Official statistics should be thorough and accompanied by appropriate metadata to ensure they are comprehendible by a wider demographic, especially those lacking specialiaed statistical knowledge[8]. As stated in the UN Fundamental Principles of Official Statistics (UNFPOS) Principle 3, "To enable accurate interpretation of the data, statistical agencies shall present

[8] https://unstats.un.org/capacity-development/handbook/chapters/Ch10_Handbook_20230417.pdf

information in line with scientific standards regarding the sources, methods, and procedures used in the statistics."

10.1. Objectives

10.1.1. Enhance accessibility:

Ensure that data is easily accessible to all stakeholders, including government agencies, researchers, policymakers, and the public.

10.1.2. Promote transparency:

Cultivate a culture of transparency by providing access to data in an open manner while upholding principles of data integrity and confidentiality.

10.1.3. Improve usability:

Deliver data in user-friendly formats that cater to the diverse needs of stakeholders, enabling them to utilise the information as per their comfort.

10.1.4. Ensure timeliness:

Disseminate data promptly to support real-time decision-making and policy formulation.

10.1.5. Standardise processes:

Establish uniform standards and procedures for data dissemination across all provincial bureaus to ensure consistency and reliability.

10.2. Mechanisms

10.2.1. Data preparation and validation:

Conduct thorough data validation checks to corroborate its accuracy and reliability while also standardising data formats and structures to facilitate seamless dissemination.

10.2.2. Data cataloguing and storage:

Maintain an updated catalogue of all available datasets, particularly metadata descriptions. Establish secure storage of data with proper backup mechanisms to prevent data loss.

10.2.3. Data release policies:

Develop and implement clear policies for data release, including access rights and confidentiality protocols. Define the frequency and schedule for data updates and releases.

10.2.4. Dissemination channels:

Utilise multiple dissemination channels, such as official websites, online portals, and public dashboards, to reach a broader and diverse audience. Leverage social media and other digital platforms to enhance outreach reach and impact of data dissemination efforts.

10.2.5. User engagement and support:

Engage with stakeholders to understand their data needs and preferences. Provide training and support to users on how to access and utilise the disseminated data effectively.

10.2.6. Feedback mechanisms:

Establish mechanisms for users to provide feedback on data quality, accessibility, and usability. Regularly review and incorporate user feedback to improve data dissemination practices.

10.2.7. Monitoring and evaluation:

Continuously assess the effectiveness of data dissemination efforts and identify areas for improvement. Conduct periodic evaluations to measure the impact of the data dissemination framework on stakeholder engagement and decision-making processes.

10.2.8. Targeted data products and services

Data products and services should be tailored to meet the needs of the diversified user groups that include:

- The general public
- Public-sector organisations
- Private-sector organisations
- Private-sector investors
- International organisations
- Specialised researchers and research centres
- Printed and audio-visual media
- Students at universities, institutes, and postgraduate programmes.
- School students of all stages

10.3. Standard Operating Procedures

10.3.1. Provincial data dissemination policy

- The PBS dissemination policy shall be referred to as the National Data Dissemination Policy, followed by both federal and provincial bureaus of statistics.
- Data dissemination policy shall be reviewed and updated no later than five years to ensure it aligns with evolving statistical requirements, technological advancements, and stakeholder needs.

10.3.2. Quality assurance for data dissemination

- Establish a quality assurance process for verifying the accuracy, consistency, and completeness of the data before dissemination. This includes validation checks, peer reviews, and cross-verification with historical data or data from other sources.

10.3.3. Adherence to the data release calendar

- Ensure strict adherence to the data release calendar to maintain transparency and build trust among data users.
- Any delays in data dissemination shall be communicated promptly with reason and revised timelines.

10.3.4. User-friendly data portal

- Maintain an online data portal that provides easy and open access to statistical data.
- Ensure the platform is user-friendly and regularly updated with the latest data releases.
- Provide relevant information, such as metadata, to ensure a complete understanding of the data being accessed.
- Employ data analytics, such as interactable figures and graphs, on portals to provide a simple understanding of key metrics.

10.3.5. Standardised data formats for data dissemination

- Provide data in multiple accessible formats, such as PDFs for reports and CSV/Excel for raw datasets, to cater to a wide range of users and their technical capabilities.
- Facilitate effective communication and engagement with stakeholders for feedback and transparency.
- Assurance of consistency of data formats over time periods when reporting data that has expanded over many years.

10.3.6. Feedback mechanism and data producer and user workshops

- Organise regular consultations with key stakeholders (e.g., policymakers, researchers, businesses) to gather feedback on data requirements to ensure that the bureau's dissemination strategies meet user requirements.
- Each PBoS shall organise *Data Producer and User Workshops* annually.

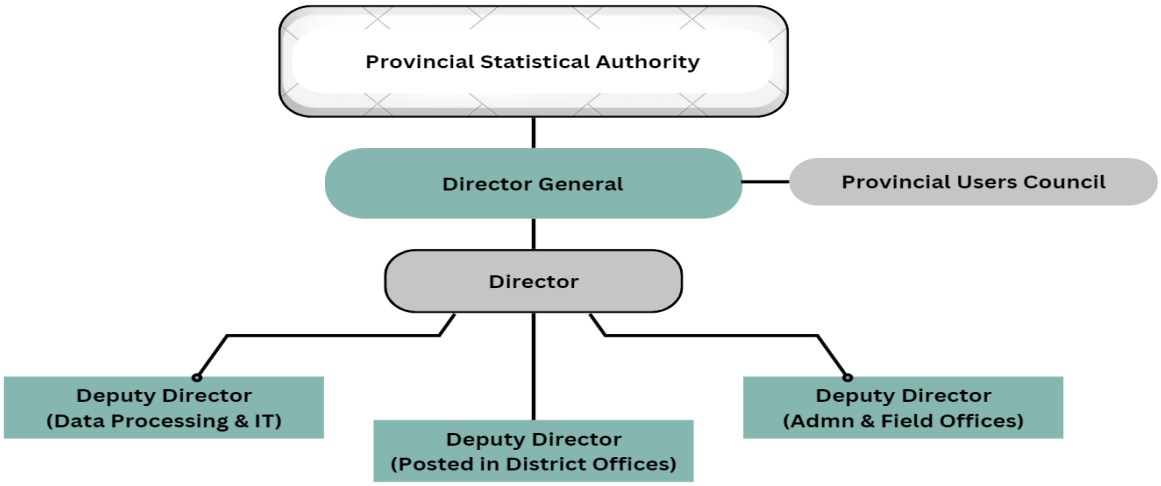
Appendix A

National Statistical System (NSS) of Pakistan

At the core of the country's NSS is the Pakistan Bureau of Statistics (PBS), which serves as the central agency or the National Statistical Office. The NSS further comprises four provincial bureaus of statistics (hereafter PBoS), one in each of the provinces, along with two statistical cells in Gilgit-Baltistan and Azad Jammu and Kashmir (AJ&K) that are yet to possess full-fledged bureaus.

The PBoS are responsible for collecting, compiling, and disseminating statistics on policy areas devolved to the provinces following the 18th Constitutional Amendment.[9] They also coordinate among the various statistical cells of provincial departments to collect secondary data and ensure cohesive data management and reporting. Figure 1 provides an organogram that outlines the hierarchical structure of the PBoS within the provincial government's framework.

Figure 1: Organogram for Provincial Bureaus of Statistics



This organogram illustrates a structured approach to data governance within the provincial statistical system, highlighting clear lines of authority and responsibility. The inclusion of specialised roles for data processing and IT underlines the importance of technological support in modern statistical operations, ensuring the PBoS can efficiently handle large volumes of data and maintain high standards of data quality and security.

Despite a clear functional hierarchy, there is no standardised framework guiding data governance at the provincial level. Provinces lack any clear direction on a mandate for data collection, rules of business for coordinating with departments to collect secondary data, a strategy for data dissemination, and any SOPs on data security. The country's NSS lacks a robust Data Governance Framework.

[9] https://www.sbp.org.pk/reports/annual/aarFY23/Chapter-07.pdf


Data for
Development


Netherlands


SDPI
Sustainable Development Policy Institute


UNFPA